

# Canonical Modeling: Review of Concepts with Emphasis on Environmental Health

Eberhard O. Voit

Department of Biometry and Epidemiology, Medical University of South Carolina, Charleston, South Carolina, USA

The article reviews concepts of canonical modeling in the context of environmental health. Based on biochemical systems theory, the canonical approach was developed over the past thirty years and applied to complex systems primarily in biochemistry and the regulation of gene expression. Canonical modeling is based on nonlinear ordinary differential equations whose right-hand sides consist of products of power-law functions. This structure results from the linearization of complex processes in logarithmic space. The canonical structure has many intriguing features. First, almost any system of smooth functions or ordinary differential equations can be recast equivalently in a canonical model, which demonstrates that the model structure is rich enough to deal with all relevant nonlinearities. Second, a large body of successful applications suggests that canonical models are often valid and accurate representations of quite complex, real-world systems. Third, a set of guidelines supports the modeler in all phases of analysis. These guidelines address model design, algebraic and numerical analysis, and the interpretation of results. Fourth, the structure of canonical models, especially those in S-system form, facilitates algebraic and numerical analyses. Of particular importance is the derivation of steady-state solutions in an explicit symbolic or numerical form, which allows further assessments of stability and robustness. The homogeneous structure of canonical models has also led to the development of very efficient, customized computer algorithms for all steps of a typical analysis. Fifth, a surprising number of models currently used in environmental health research are special cases of canonical models. The traditional models are thus subsumed in one modeling framework, which offers new avenues of analysis and interpretation. **Key words:** allometry, biochemical systems theory, canonical model, Cox model, exposure, extrapolation, risk assessment, systems analysis, S-system. — *Environ Health Perspect* 108(suppl 5):895–909 (2000).

<http://ehpnet1.niehs.nih.gov/docs/2000/suppl-5/895-909voit/abstract.html>

with health status, and effects can be complicated by the multiple exposures that are typical of the human experience.

The present review focuses exclusively on issues of modeling. In this context, an integrative approach, as suggested by Olden and Klein, requires novel strategies of merging various types of information from the different branches and hierarchical levels of environmental systems into comprehensive yet manageable mathematical structures. To this end, the review proposes to study the applicability of a methodologic framework, called canonical modeling, to integrated environmental health analyses. Canonical modeling was originally designed to deal with biochemical systems and was called biochemical systems theory but later found applications far beyond biochemistry. The objective of this approach, independent of the specific application area, is to capture the dynamics of complex systems that are characterized by numerous components and large numbers of mostly nonlinear interactions.

Because of their nonlinearities, complex systems defy the law of superposition [which Garfinkel (2) traces back to Julius Caesar's motto *Divide et Impera*, or *Divide and Conquer*]. Linear systems can be disassembled, their parts can be studied in isolation, and the subsequent reassembly of individual responses into an overall response is mathematically valid. In contrast, nonlinear systems usually lose essential characteristics when taken apart. An organism or ecosystem is much more than a disorganized mix of chemicals. In nonlinear systems, superposition is supplanted by synergisms and antagonisms; overall responses are stronger, weaker, or entirely different than the sum of all individual responses. This difference between linear and nonlinear systems is of crucial importance because all simplifications resulting from studying one component at a time are no longer automatically justified. In fact, they are often invalid.

## Introduction

Questions of environmental health confront us with the most complex phenomena science has yet attempted to analyze. Environmental systems have geologic, physical, chemical, and biologic components, and these are intricately connected to human health, behavior, and policy decisions. Over the past decades, the attempt to understand these connections has led to a new scientific discipline: quantitative risk assessment. Given the anthropocentric attitude of society, aspects of human health originally received priority among the risk assessment activities, but with the realization of the intimate connections between human and environmental health, interest in ecologic systems and the environment followed quickly.

Even though human and environmental risk assessments have much in common, their different emphases have resulted in quite different philosophic and mathematical approaches. In human risk assessment, the individual receives attention, and potentially fatal hazards are considered along with comparably less significant consequences such as rashes or headaches. In contrast, ecologic risk assessment is characterized by large numbers of individuals and species and their interactions, and by questions of ecosystem stability and diversity. Ecologic assessments often target the survival of entire populations or the sustainability of an ecosystem, and individual

deaths or illnesses are rather readily tolerated. Even the local extinction of populations is often accepted unless these populations belong to globally endangered species. Changes in ecosystem composition are monitored but are not necessarily cause for concern per se.

Ecologic risk assessors have made great strides in characterizing the effects of contaminants on integrated environmental systems. By contrast, human health risk assessors have traditionally tried to associate a single disease with a single cause, thereby ignoring the connectivity of environmental systems. Although the differences in goals and methods of human and environmental risk (or health) assessment are understandable, it appears that an integrative approach to assessing environmental hazards with respect to human and environmental health would be better balanced and more cost effective than two separate frameworks of analysis. As Olden and Klein (1) put it:

Along with understanding the basic building blocks of cellular and molecular systems and their interactions with environmental agents, we must integrate these single-system studies into multi-system studies. This effort is especially important for the study of environmental effects because environmental agents can affect different organs differently, effects can vary over time and

This article is part of the monograph on Mathematical Modeling in Environmental Health Studies.

Address correspondence to E.O. Voit, Dept. of Biometry and Epidemiology, Medical University of South Carolina, PO Box 250551, 1148 Rutledge Tower, 135 Rutledge Ave., Charleston, SC 29425 USA. Telephone: (843) 876-1122. Fax: (843) 876-1126. E-mail: [voiteo@musc.edu](mailto:voiteo@musc.edu)

Received 15 February 2000; accepted 28 July 2000.

Before I review methodologic advances in addressing complex systems, it might be useful to identify what exactly it is that makes a system complex. The first feature is a large number of components. In terms of environmental issues, one might think of the thousands of organisms making up the rainforest, but one doesn't have to go that far. An organism like a yeast cell is rather simple in comparison to a higher animal or plant, yet it consists of tens of thousands of biochemical components. The human brain contains on the order of 100 billion neuronal components and hundreds of trillions of interconnections (3). The human body is composed of some five octillion atoms (4). Just describing, not even analyzing, these components poses an enormous bookkeeping problem that can only be managed by mathematical means. These magnitudes are the rule rather than the exception, leading Goldenfeld and Kadanoff (5) to remark almost cynically, "Everything is simple and neat—except, of course, the world." The large numbers of components have to be contrasted with the capacity of the human mind. Supported by an array of psychologic experiments, Miller (6) suggested that the human mind is limited to simultaneously processing  $7 \pm 2$  pieces of information. Clearly, that capacity unaided does not go very far.

The second crucial feature of complex systems is the ubiquitous hierarchy of processes. Processes occur in large numbers at different levels of organization and at different temporal scales. The environment is full of examples. One may only think of the control of population size, which at one level is determined by growth rates, mortality, predation, immigration, and other high-level processes, but at a lower level also is the consequence of processes at the biochemical and physiologic levels of individuals.

Adding to the complications caused by their large numbers and hierarchical organization, most processes governing complex systems are nonlinear. Beyond the invalidity of the principle of superposition, nonlinearities cause two problems for any intuitive, non-mathematical approach. First, they create situations that lead our thinking in terms of causes and effects astray. Without the aid of mathematical analysis, one cannot even reliably predict the effects of the simplest regulatory mechanisms such as the ubiquitous feedback inhibition. If an input to the system is increased, does the output decrease (because of the inhibition)? Does it increase (in spite of the feedback)? Is it unaffected? Questions of this nature cannot be answered with intuition alone. In fact, one can show that qualitatively different responses in a feedback loop are possible, depending on the numerical specifications of the system (7).

In more general terms, nonlinearities make it difficult to find intuitive explanations for why systems in nature are designed the way they are, and it is often impossible reliably to predict the correct system response without a rigorous quantitative approach (8). For instance, it is impossible to evaluate the advantages of inhibition of product formation over activation of its degradation, even though both can generate the same overall system responses (7–9).

The second problem caused by nonlinearities is that the qualitative types of system responses may depend on quantitative properties of the system: If some parameter is within a certain range, the system hovers around a normal state; if it is less than some lower threshold value, the system may exhibit sustained oscillations; and when the parameter exceeds some higher threshold, the system may cease to function altogether. With numerous parameters having these characteristics, there is no possibility to understand or predict the effects of changes in some system components without a mathematical analysis (8). While the "lure of the linear" (10), with its solid theoretical underpinnings and a great repertoire of powerful tools, has great appeal for the mathematician, nature simply is not linear. It is therefore necessary for the scientific community to develop germane methods of capturing the essence of nonlinear phenomena.

The nonlinear character of environmental systems and the associated invalidity of the principle of superposition constitute a true challenge. The main philosophy of science over the past hundred years has been reductionism, and since this approach implicitly assumes superposition, it is no longer applicable if the focus of an investigation is the integrated nature of a system. As Savageau (11) pointed out:

Paradoxically, it is at the very height of its success that the weaknesses of this paradigm [of reductionism] are becoming apparent. We shall soon have the complete parts catalog of *E. coli*. Yet, by comparison, we still know little about the integrated system, what makes it a living cell, or how it will respond to novel environments, and to specific changes in its molecular constitution. Our knowledge is fragmented and descriptive; we have almost no understanding of the "design principles" that govern the intact biological system.... We need a radically different approach that is able to elucidate quantitative and qualitative features of complex integrated systems.

Summarizing the state of the art of reductionism, Simpson (12) concluded that reduction to atomic and molecular levels alone is neither philosophically nor practically sufficient and that all levels of the biologic hierarchy have to be studied if

biological phenomena are to be explained. Yates (13), a strong supporter of studying complexity and a proponent of an integrated approach to biology, found that even among true believers in reductionism "there is a residual mystery after the reduction is accomplished."

The failure of reductionism to yield a true understanding of phenomena in nature has led to the advent of a paradigm shift (14) in scientific thinking. It has become evident that new laws or reconstruction have to be postulated—not laws describing parts, but laws capturing the responses of integrated systems (11). With a paradigm shift from reduction to reconstruction unfolding before us, one must ask what the new challenges are and what types of strategies one must devise to address them. Three components seem to be essential for accepting the challenge: a valid means of problem simplification, a convenient terminology, and a convenient mathematical representation for large systems.

## Simplification

Nature itself has afforded us with different types of simplifications for the analysis of complex systems. They are based on organizational, temporal, and spatial hierarchies (7). The typical organizational simplification consists of intentionally ignoring very much lower or very much higher levels of biologic organization. If the relationships between an environmental contaminant and the prevalence of a disease are the focus of investigation, the laws of particle physics are certainly still in effect, but it is unnecessary and undesirable to represent the disease mechanisms in these terms. Thus, one simplification consists of replacing the complexity at all molecular and submolecular levels with some average behavior. By the same token, one is often justified to ignore processes at higher levels, such as evolution and long-term changes in climate, when processes at a lower level are the focus.

In many instances, the hierarchy in organizational levels corresponds to a hierarchy of time scales at which processes occur. In an environmental context, one could, for instance, differentiate quantum physical, biochemical, physiologic, developmental (with respect to individuals and populations), evolutionary, and geologic time scales. Limiting an analysis to a single time scale provides a significant simplification, as processes occurring at much faster speeds approach their steady states so fast that their dynamics are irrelevant, whereas processes with much slower rates don't change much and thus are essentially constant (7,15).

Another simplification results from the spatial organization of natural systems. Species occupy niches, forming a "patchy" environment. This spatial heterogeneity

strongly affects the patterns of interactions between species. For instance, diseases do not move randomly throughout an ecosystem, especially if it is as big as the globe, but often spread in patterns that resemble directed diffusion processes or traffic patterns. At a cellular scale, enzymes do not randomly collide with all kinds of proteins until they incidentally find their substrate. They are confined to compartments, attached to channels or surfaces, or exist in complexes that service consecutive steps of metabolic pathways (16–19). These and other natural simplifications translate directly into significant mathematical simplifications that can and should be exploited in approaches to analyzing complex systems (7).

### Terminology

The second ingredient of an integrated systems approach is an efficient and convenient terminology. Selecting terminology is not just a cosmetic decision but a crucial step in the modeling process because it determines what types of questions can be asked and answered by the analysis. The terminology of canonical modeling is a subset of applied mathematics. One distinguishes dependent variables, independent variables, and parameters, and describes the dynamics of the system by formulating differential equations that describe how each dependent variable changes over time.

A dependent variable represents a system component or a pool of components whose numerical value (e.g., its size or concentration) is affected by the system and the environment. Typically, the values of dependent variables change during an experiment. In an age-structured population model, the dependent variables are sizes of subpopulations with a given age. In a toxicologic model, they may represent biologic reactive intermediates.

An independent variable represents a system component, or a pool of components, that is unaffected by the system. In a population model, the number of immigrants of a given age may be an independent variable. In a toxicologic model, the chemical agent of concern may be coded as an independent variable. Typically, independent variables are constant during any given mathematical experiment or they change in a manner that is controlled by the experimentalist or the environment but not by the investigated system.

A parameter quantifies some property of the system and is defined as an entity with a constant numerical value. A parameter could be the size of the territory occupied by the population or a buffered pH in an *in vitro* experiment.

### Mathematical Representation

The standard approach to describing the dynamics of systems is to study the temporal

changes in all components. To this end, one formulates for each dependent variable a differential equation, whose right-hand side has a valid and convenient form. The quality of this valid and convenient form is measured against several requirements. Ideally, the mathematical representation (model) should satisfy the following criteria:

- It should capture the essence of the system under realistic conditions. The model should respond to relevant inputs the same way as the actual system. Ideally, the model would respond correctly under all imaginable conditions, but this would be too restrictive a requirement. Instead, one requires the model to react correctly under most realistic conditions. If a system *in situ* is being investigated, then *in situ* conditions should determine what realistic means. This relaxation of performance criteria has significant consequences. For instance, one may create experimental conditions in the lab that span several orders of magnitude in concentrations of a contaminant and study the responses of select organisms. If the contaminant in reality only occurs in a limited concentration range, a mathematical model describing the effects of the contaminant in the environment should be required to capture the responses in this limited range but not necessarily outside this range.
- It should be qualitatively and quantitatively consistent with key observations. The model should offer one-to-one relationships between actual observations and the mathematical representation. A good model should render it possible to execute mathematical experiments that correspond, one to one, to those executable in the environmental system of interest.
- It should, in principle, allow analyses of arbitrarily large systems. The structure of environmental models should be independent of the size of the system. An example of a size-independent structure is a system of linear equations: The addition of a new variable increases the number of terms in each equation and requires a further equation, but the system is still linear, its structure is unchanged, and all methods of linear analysis still apply. In contrast, if one adds an inhibitor to a traditional toxicologic model, the structure of several rate laws is likely to change, and there are no clear recipes for implementing these changes.
- It should be generally applicable. Many models require the investigator to know or assume the precise mathematical form of all processes of a system before the model can be analyzed. This requirement becomes very severe in environmental

systems. A preferable modeling approach must allow the investigator to formulate model equations directly from the structure of the analyzed phenomenon, i.e., from a diagram that shows the flow of material and the existence of modulations.

- It should be characterized by measurable quantities. Ideally, every variable and every parameter should have a uniquely defined, measurable role and meaning. A system component may represent a feature that is not traditionally measured or cannot be measured with today's methods, but it should, in principle, be a measurable feature.
- It should allow simple translation of results back into subject area language, be it toxicology, ecology, epidemiology, or health risk assessment. Once a mathematical model is formulated, its analysis consists of procedures of applied mathematics and computer simulation, and typical results are expressed in mathematical terminology. A good modeling approach should render it easy to translate what the numeric or symbolic results mean in terms of environmental health assessment.
- It should have a mathematical form that is amenable to analytic and numeric evaluation. All the previous features were concerned with the appropriateness of a good mathematical representation. Almost as important is its mathematical tractability. There is little use for a mathematical model, no matter how appropriate, if one cannot analyze it.

No mathematical model is known to combine all desired features in an ideal fashion. In many cases, a model either is mathematically very complex and accurately represents some set of observations or it allows standard mathematical analyses but shows deficiencies in comparison with the actual system. The multitude and complexity of processes governing environmental phenomena almost always preclude the direct use of traditional mechanistic models that deal with isolated processes. For instance, if one asks for the quantitative relationship between some environmental contaminant and the ultimate health effects in humans, it is obvious that no simple algebraic function is able to describe this relationship.

The only feasible alternative lies in the prudent use of approximations. Accompanying this switch from detailed mechanisms to approximations, the types of questions to be asked shift from individual processes toward a more global cooperation of system components. When analyzing integrated system models, one may ask how the temporal change in one variable affects the concentration of the other variables, how the system is affected by changes in modulation, or why nature has

selected a particular mode of regulation over other possible modes. One may ask what types of perturbations an organism can tolerate or which components of a system are most sensitive to such perturbations.

Among the best compromises currently known are models in which the interactions between components are represented as products of power functions. These functions may appear to be rather unusual, but they are mathematically and logistically justifiable as well as convenient.

## Canonical Modeling

The canonical modeling approach is based on key papers of Savageau (20–22) and was subsequently expanded in collaboration with several laboratories around the world. In most cases, this methodology has been applied to questions in biochemistry and the regulation of gene expression. It is adapted here to questions of environmental health assessment. Several hundred articles and book chapters about this methodology have appeared in the literature. For summaries relevant in the present context see, for instance (23–27).

The formulation of dynamic models for complex systems begins with general equations that describe how each system component changes over a period of time. Although it is usually impossible to formulate directly the value of a component as an explicit function of time, one can often determine from observations which processes contribute to the increase or augmentation of a component and which contribute to its decrease or depletion. If this information is translated into mathematical terminology, the result is a system of differential equations.

For streamlined notation, suppose the term  $X_i(t)$  describes the status of a dependent variable, which is one that actually changes during the time period of interest, at a time point  $t$ . For  $n$  components of this type ( $i = 1, \dots, n$ ) one may thus symbolically assert

$$\dot{X}_i(t) = V_i[X_1(t), X_2(t), X_3(t), \dots, X_n(t)] \quad i = 1, 2, \dots, n. \quad [1]$$

The functions  $V_i$  may be extremely complicated, but if we assume for a moment that we knew these functions and the status of the system at the beginning of our observation period, then the equation contains all possible information about the system. Of course, in reality the functions are not known. Nonetheless, this starting point is useful, as it focuses the problem of capturing the overall dynamics of a large system to the mathematical characterization of clearly localized processes.

## S-Systems

It is useful to amend Equation 1 slightly by splitting up the functions  $V_i$  into two

functions each, reflecting that the change in a component is the balance between all processes of production or augmentation of  $X_i$  and those of degradation or depletion. The differential equations accounting for the combined dynamics read

$$\begin{aligned} \dot{X}_i &= V_i^+(X_1, X_2, X_3, \dots, X_n) \\ &\quad - V_i^-(X_1, X_2, X_3, \dots, X_n) \\ i &= 1, 2, \dots, n. \end{aligned} \quad [2]$$

As an alternative, one could argue that many processes may contribute to the production and to the depletion of  $X_i$  and that, therefore, the right-hand side of the differential equation should consist of many functions. Mathematically speaking, this scenario is already included in Equation 2, as  $V_i^+$  and  $V_i^-$  are not yet specified and can, in particular, be sums of functions that each represents one individual reaction. However, there is a semantic difference between these formulations, which leads to alternate representations, as will be discussed later.

In addition to the dependent variables, whose values change over time, it is useful to include independent variables in the model formulation. These variables are constant for the duration of each experiment but may change from one experiment to the next. Independent variables are not formulated as differential equations, as they are constant, and the right-hand sides of their differential equations would equal zero. Nevertheless, the independent variables are included in right-hand sides of the system equations to make their influence explicit. Accounting for  $n$  dependent and  $m$  independent variables, the general differential equations for the entire system are thus

$$\begin{aligned} \dot{X}_i &= V_i^+(X_1, X_2, \dots, X_n, X_{n+1}, X_{n+m}) \\ &\quad - V_i^-(X_1, X_2, \dots, X_n, X_{n+1}, X_{n+m}) \\ i &= 1, 2, \dots, n \end{aligned} \quad [3]$$

These  $n$  differential equations are called the system equations, or, collectively, the system equation, or the system description. Sometimes they are also referred to as (Kirchhoff's) node equations because they show some similarity to equations describing the flow of currents in branched electric circuits (28). When one talks about the behavior or response of a system, one means the collective change of all its constituents.

In the mathematical formulations above, the system descriptions are very general, as the functions  $V_i^+$  and  $V_i^-$  on the right-hand sides have not been specified. Identification of appropriate functions is the key to a successful systems analysis. As discussed above, there is no hope of finding explicit, mechanistic functions that could be used here, and because biologic and environmental systems are essentially always nonlinear, one comes to the con-

clusion that  $V_i^+$  and  $V_i^-$  are to be found as convenient nonlinear approximations.

In most situations, nothing really definite is known about the actual processes that are to be captured by the functions  $V_i^+$  and  $V_i^-$ , but for simplicity of argument it is reasonable to assume that they are sufficiently smooth (differentiable; without gaps or sharp corners) and positive-valued. If the latter is not the case, the problem can be reformulated in terms of additional functions to ensure positivity (29).

The crucial concept of canonical modeling is a suitable representation of  $V_i^+$  and  $V_i^-$  in terms of power-law functions. The derivation of this representation follows rigorous, well-established methods of numerical analysis and has been discussed numerous times (7,20,23,27). In a nutshell, the functions and variables are represented in logarithmic coordinates. In this coordinate system, the functions are approximated by Taylor series, where only the constant and linear terms are retained. In other words, the functions are linearized in logarithmic space. The linearized functions are subsequently translated back into Cartesian coordinates. The result of this strategy is a representation of  $V_i^+$  and  $V_i^-$  as products of power-law functions in the dependent and independent variables:

$$\begin{aligned} V_i^+ &\approx \alpha_i \prod_{j=1}^{n+m} X_j^{g_{ij}} & i = 1, 2, \dots, n \\ V_i^- &\approx \beta_i \prod_{j=1}^{n+m} X_j^{h_{ij}} & i = 1, 2, \dots, n \end{aligned} \quad [4]$$

The multipliers  $\alpha_i$  and  $\beta_i$  in these terms are rate constants that characterize the flux rates between pools or variables. The exponents  $g_{ij}$  and  $h_{ij}$  are called kinetic orders. Their numerical values reflect the strengths of the effects that the corresponding variables have on a given flux term. A large positive value signifies a strong activating or augmenting effect, a negative value signifies inhibition, and a value close to 0 indicates that the corresponding variable is not very influential in the given flux term. Although the representations in Equation 4 formally include all dependent and independent variables, it is not difficult to see that ultimately only those variables are included that directly affect the functions  $V_i^+$  and  $V_i^-$ . All other variables have exponents of 0, which effectively eliminates them from the corresponding terms.

To appreciate the interpretation of the exponents  $g_{ij}$  and  $h_{ij}$  as kinetic orders, recall the basics of elemental chemical kinetics. In a bimolecular reaction of the type  $X_1 + X_2 \Rightarrow X_3$ , the production of  $X_3$  is equal to the product of a rate constant,  $k$ , and the two concentrations of  $X_1$  and  $X_2$ :

$$\text{Production of } X_3 = \dot{X}_3 = kX_1X_2. \quad [5]$$

If the reaction involves two molecules of  $X_1$  and one molecule of  $X_2$ , then  $X_1$  enters the change equation with a power of 2,

$$\dot{X}_3 = kX_1^2X_2, \quad [6]$$

and one says that the kinetic order of the reaction with respect to the chemical species  $X_1$  is 2 and that the kinetic order of the reaction with respect to the chemical species  $X_2$  is 1. In elemental chemical kinetics, these powers are directly interpreted as the numbers of molecules involved in each individual reaction, and as a consequence, these powers assume values like 1 and 2. In canonical modeling, they are allowed to assume any real values. It is noted that the same type of formulation of associate processes as products of variables is also found in the famous predator-prey models of Lotka (30) and Volterra (31) and their generalizations.

At one point of choice, the operating point, the representations in Equation 4 are exactly equivalent to the original functions  $V_i^+$  and  $V_i^-$  if the numerical values of the rate constants and kinetic orders are chosen appropriately. In many cases (but not necessarily), the operating point is chosen as a steady-state point, which is a point at which the system, in an overall sense, does not change. This point often consists of the nominal values of all variables, i.e., of values that are found under normal (steady-state) conditions *in situ*. Close to the operating point, the power-law representations are guaranteed by Taylor's theory to be very accurate. Ample experience has demonstrated that the range of valid representation is often wide, sometimes spanning several orders of magnitude in the values of the dependent and independent variables (7,23,32).

If the power-law approximations (Equation 4) are substituted in the general system description (Equation 3), the result is a highly structured system of differential equations known in the biomathematical literature as an S-system:

$$\dot{X}_i = \alpha_i \prod_{j=1}^{n+m} X_j^{g_{ij}} - \beta_i \prod_{j=1}^{n+m} X_j^{h_{ij}} \quad i = 1, 2, \dots, n. \quad [7]$$

It is noted that this functional form is always the result, when the functions  $V_i^+$  and  $V_i^-$  in Equation 3 are linearized in logarithmic coordinates, independent of their mathematical structure. The only pieces of information used in this approximation are the operating point and the slopes of  $V_i^+$  and  $V_i^-$  at this point in logarithmic coordinates. Just as smooth functions of one or two arguments can be approximated with a straight line or a plane, a realistic system can always be approximated by simple power-law functions or their multivariate products.

A very important consequence is that the approximation strategy is applicable even if the functions  $V_i^+$  and  $V_i^-$  are unknown. This may be surprising at first glance but again is analogous to the fact that any unknown functions, as long as they are smooth (differentiable), can be locally represented by linear functions. If the true functions  $V_i^+$  and  $V_i^-$  are unknown, the numerical values of the parameters in the power-law representation (Equations 4 and 7) are not known, but the structure of the equations is always the same. Consequently, this structure can be formulated without knowledge of the original functions and is based solely on information about which variables directly affect  $V_i^+$  and  $V_i^-$ .

### Examples

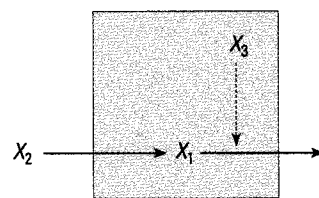
**Example 1.** Consider the system in Figure 1 with only one dependent and two independent variables. In this system,  $X_2$  is supplied by the experimenter at a constant rate. The independent component  $X_2$  is converted into  $X_1$ , which is subsequently degraded. The degradation is inhibited by  $X_3$ . Even though  $X_3$  is inside the system, it is considered an independent variable because the dynamics of  $X_1$  does not change the value of  $X_3$ .  $X_3$  is simply sending a signal, but there is no flow of material.

Because there is only one dependent variable, the S-system contains only one equation. The mathematical description in symbolic form thus reads

$$\begin{aligned} \dot{X}_1 &= \alpha_1 X_2^{g_{12}} - \beta_1 X_1^{h_{11}} X_3^{h_{13}} & X_1(t_0) &= X_{10} \\ X_2 &= \text{const.} \\ X_3 &= \text{const.} \end{aligned} \quad [8]$$

Without further information, the values of the parameters and the (constant) independent variables  $X_2$  and  $X_3$  are unknown. It is known, though, that only  $h_{13}$  is negative, as it represents an inhibitory effect. The remaining kinetic orders  $g_{12}$  and  $h_{11}$ , as well as the rate constants  $\alpha_1$  and  $\beta_1$ , are positive.

**Example 2.** This example may be considered an expansion of the previous example in which  $X_2$  is now a dependent variable whose production is also under investigation. Specifically,  $X_2$  is the product of a reaction



**Figure 1.** An externally supplied substrate  $X_2$  is converted into  $X_1$  and subsequently degraded.  $X_3$  inhibits the degradation.

that uses  $X_4$  as substrate and is activated by  $X_1$  and inhibited by  $X_3$ . A graphical representation is given in Figure 2. The single equation of the previous example is augmented with an equation describing the dynamics of  $X_2$ . The symbolical system description is

$$\begin{aligned} \dot{X} &= \alpha_1 X_2^{g_{12}} - \beta_1 X_1^{h_{11}} X_3^{h_{13}} \\ X_1(t_0) &= X_{10} \\ \dot{X}_2 &= \alpha_2 X_1^{g_{21}} X_3^{g_{23}} X_4^{g_{24}} - \beta_2 X_2^{h_{22}} \\ X_2(t_0) &= X_{20} \\ X_3 &= \text{const.} \\ X_4 &= \text{const.} \end{aligned} \quad [9]$$

Again, one deduces from the structure of the system that  $h_{13}$  and  $g_{12}$  are negative, as they represent inhibitory effects; all other parameters are positive. Furthermore, the numerical values of  $g_{12}$  and  $h_{22}$  are equal, as are the values of  $\alpha_1$  and  $\beta_2$ , because the degradation of  $X_2$  and the production of  $X_1$  are identical.

With typical numerical specifications chosen for illustrative purposes, the system may read

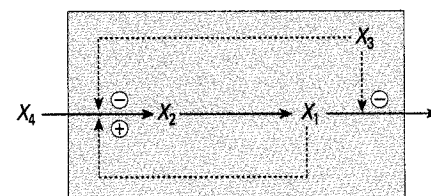
$$\begin{aligned} \dot{X}_1 &= 2X_2 - 1.2X_1^{0.5}X_3^{-1} & X_1(0) &= 2 \\ \dot{X}_2 &= 2X_1^{0.1}X_3^{-1}X_4^{1/2} - 2X_2 & X_2(0) &= 0.1 \\ X_3 &= 0.5 \\ X_4 &= 1. \end{aligned} \quad [10]$$

The dynamics of the system is shown in Figure 3. It is obtained by numerically solving the differential equations in [10] with a general all-purpose solver like a Runge-Kutta or Gear algorithm or with specific algorithms for canonical models [e.g., PLAS (<http://correio.cc.fc.ul.pt/~aenflplas.html>); see also Voit (27)].

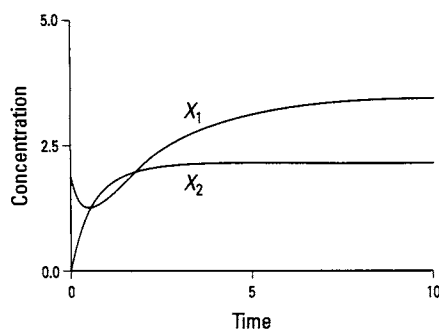
### Alternate Canonical Models

If the model is set up in a fashion where the right-hand sides contain several functions and not just one  $V_i^+$  and one  $V_i^-$ , the same type of power-law approximation may be applied to each function. For instance, the  $i$ th equation of the general system description could be

$$\begin{aligned} \dot{X}_i &= V_{i1}^+ + V_{i2}^+ + V_{i3}^+ - V_{i1}^- - V_{i2}^- \\ &\quad - V_{i3}^- - V_{i4}^- \quad i = 1, 2, \dots, n \end{aligned} \quad [11]$$



**Figure 2.** The diagram of Figure 1 is expanded to include  $X_2$  as a dependent variable whose production is modulated by  $X_3$  and an additional independent variable,  $X_4$ .



**Figure 3.** Dynamics of  $X_1$  and  $X_2$  in the example depicted in Figure 2 and formulated in Equation 10.  $X_1$  initially undershoots but ultimately reaches a level higher than at the beginning of the experiment.  $X_2$  shows a simple monotonic increase.

where each function  $V_{ij}^+$  or  $V_{ij}^-$  may depend on several or all of the dependent and independent variables  $X_1, \dots, X_{n+m}$ . In general, the result of this strategy is a collection of products of power-law functions on each of the right-hand sides. This type of canonical model is called a Generalized Mass Action (GMA) system and has the form

$$\dot{X}_i = \gamma_{i1} \prod_{j=1}^{n+m} X_j^{\varepsilon_{1ij}} \pm \gamma_{i2} \prod_{j=1}^{n+m} X_j^{\varepsilon_{2ij}} \pm \dots \pm \gamma_{ik} \prod_{j=1}^{n+m} X_j^{\varepsilon_{kij}} \quad [12]$$

where the positive coefficients  $\gamma$  are rate constants of the associated processes and the exponents  $g$  are again kinetic orders.

This type of equation, which was also called a multinomial system (33), is a legitimate system description, and many comparisons between this representation and corresponding S-systems have been performed in the context of metabolic pathway analysis. These comparisons have elucidated derivation, development, analytic and computational tractability (34,35), biochemical validity (36–38), accuracy (32), feasibility for optimization (39), and general mathematical properties (23,29,40) of the two representations.

Both representations, S-systems and GMA systems, have advantages in their own right, and depending on the purpose of the analysis, one or the other power-law representation may be preferable. Whether one or the other model is closer to the true metabolic dynamics is an unanswered question, and one can easily imagine that not one model would always be superior. Furthermore, for analyses, numeric or algebraic, that remain in close vicinity of a normal operating point, the two modeling strategies usually yield rather similar results. Nonetheless, the S-system form has great analytic advantages that derive from the fact that S-systems have linear steady-state equations; some examples will be discussed in a later section.

As a second alternative, one could argue that the difference between  $V_i^+$  and  $V_i^-$  itself is a function  $V_i$  that could be approximated by a single product of power-law functions. The general system description in this case would simply be

$$\dot{X}_i = V_i(X_1, \dots, X_{n+m}). \quad [13]$$

Again developing the Taylor polynomial, one obtains the change in  $X_i$  as

$$\dot{X}_i = \pm \gamma_i \prod_{j=1}^{n+m} X_j^{f_{ij}} \quad i = 1, 2, \dots, n \quad [14]$$

Systems consisting of these types of equations are known as Riccati systems (33) or half-systems (29). While mathematically interesting, they are generally inconvenient for the analysis of real-world applications.

### Properties of S-System Models

**Validity.** S-system models are consistent with a number of specific features of environmental systems at all levels. At the chemical level, S-system models are a direct generalization of the traditional laws of elemental chemical kinetics. At the biochemical level, power functions, as they are the basis for S-systems, appear to provide rate laws of sufficient accuracy [see Voit and Savageau (32) and Savageau (41) for a discussion of this question]. Furthermore, S-systems are consistent with modern observations that biochemical reactions in heterogeneous media show fractal kinetic orders (41–43).

At a higher level of organization, S-systems conform very well with observations of allometry, which have been made in a variety of environmental contexts. The paradigm example is allometric growth: one finds that the absolute growth of one part of an organism (or other system) is not linearly related to the growth of another part, but instead, that the relative growth of two parts very often is linearly related. If the first part grows by 5% over a certain period of time, then the other part grows by approximately the same percentage. Expressed mathematically, the relative growth of the two parts  $X$  and  $Y$  can be formulated as

$$\frac{1}{Y} \frac{dY}{dt} = c_1 \cdot \frac{1}{X} \frac{dX}{dt}. \quad [15]$$

Integration yields

$$\ln(Y) = c_1 \cdot \ln(X) + c_2, \quad [16]$$

and exponentiation produces the allometric relationship in its typical form

$$Y = \alpha X^g, \quad [17]$$

where  $\alpha = \exp(c_2)$  and  $g = c_1$ .

Allometric relationships are mathematically very convenient. On one hand, they are nonlinear and tend to describe biologic phenomena with greater accuracy than linear functions. On the other hand, these relationships are linear when represented in logarithmic coordinates, as shown in Equation 16. There are literally hundreds of examples of allometric relationships (shown in logarithmic coordinates) for different species; they include blood clearance, respiratory activity, and cardiac cycles, which are all allometrically related to body mass (44,45). A directly relevant example in the context of environmental health is the relationship between LD<sub>50</sub> and the time to metabolize one body mass equivalent of O<sub>2</sub>, which follows a power-law function that many species obey (46).

Toxicology and risk assessment regularly use allometric models for extrapolations between species and for predicting biologic responses in humans based on the corresponding responses in a laboratory animal. This strategy makes the explicit or implicit assumption that the same mechanisms are at work in different species, but that their magnitude or speed is related to other physiologic parameters that characterize the species. It would be of great benefit to test the reliability of this scaling procedure with respect to chemical susceptibility and disease development among different animal species and among different groups of animals within the same species. These comparisons would provide “experimental confidence intervals,” which would show to what degree risk scaling is appropriate in human or wildlife health assessments.

The most frequently used baseline parameter for scaling is body weight. It is known that smaller organisms have higher metabolic and physiologic rates, and in very many cases, body weight and these rates are rather accurately related by allometric relationships.

An allometric example at a different scale is provided by quantitative structure-activity relationships (QSARs). QSARs have the mathematical form of power-laws and are used to predict rate constants of unknown chemical reactions from corresponding known reactions or known reactions in structurally similar compounds (47). QSARs are of vital importance in predicting the toxicity of unknown compounds and of mixtures of contaminants, which may act synergistically or antagonistically (48).

Even higher-order characteristics often are related allometrically. For instance, the relationship between the sizes and numbers of trees in a plot very often follows the 3/2 rule, which corresponds to a power-law function with power 3/2 (49,50). Similarly, the number of wildlife species appears to be allometrically related to the inhabited area, even

though the scatter in this case is often considerable (51).

An allometric relationship is not necessarily a function of only one variable. For instance, in the context of risk assessment and interspecies scaling, Hayton (52) summarized findings showing that the systemic clearance of chemicals is a power-law function of both body weight and brain weight. Suter (53) discusses dose-response models that are power-law functions in both dose and duration.

Expanding the terminology of Equations 16 and 17, a two-variable scaling function reads

$$Y = \alpha X_1^{g_1} X_2^{g_2} \quad [18]$$

or, equivalently,

$$\ln(Y) = \ln(\alpha) + g_1 \cdot \ln(X_1) + g_2 \cdot \ln(X_2). \quad [19]$$

For more variables,  $Y$  takes the form of  $V_i^+$  or  $V_i^-$  in Equation 4.

Savageau (54) showed that S-system models of growing organisms under rather unrestricted conditions exhibit this type of allometry. Voit (55) demonstrated that S-system models are unique in the sense that all their components show allometric dynamics.

It is becoming increasingly evident that environmental assessments depend on appropriate time scales. This is particularly relevant in ecologic assessments, where organisms with drastically different life expectancies are exposed to the same hazards but respond with widely differing strengths (56,57). One thus has to ask whether time scaling is compatible to canonical models. The answer is that allometric time scaling does not alter the structure of canonical models. If one makes the original time scale in an S-system model explicit by calling it  $X_0$  and formally including it on the right-hand sides, one obtains

$$\begin{aligned} \dot{X}_0 &= 1 \\ \dot{X}_i &= \alpha_i \prod_{j=0}^n X_j^{g_{ij}} - \beta_i \prod_{j=0}^n X_j^{h_{ij}}. \end{aligned} \quad [20]$$

The first equation in this system results from the fact that  $\dot{X}_0 = dX_0/dt = dt/dt = 1$ . Introduction of a new time scale  $\tau = X_0^p$ , application of the chain rule, and possibly renaming of the new time variable returns exactly the same mathematical form. In the simple case of proportional scaling (57), all  $\alpha$  and  $\beta$  coefficients are simply multiplied with the same scaling factor, whereas the exponents are unchanged.

The natural scalability of canonical power-law models must be seen in contrast to alternative models. For example, consider a basic Michaelis-Menten model of an enzyme-catalyzed reaction. If the substrate is replaced

with a power-law function  $S = \alpha R^g$ , the resulting rate law is

$$v = \frac{\tilde{V}_{max} R^g}{K_M + R^g}, \quad [21]$$

which has the form of a Hill equation. This form is qualitatively different from a Michaelis-Menten rate law in that it is s-shaped and has a 0 slope at 0 concentration, whereas the Michaelis-Menten law is simply saturated and has a slope of 1 for 0 substrate.

**Telescopic property.** An intriguing feature of the S-system form is its so-called telescopic property (58,59). It is often of interest to model a phenomenon at different levels. In a first analysis, one could be interested in the enzyme-catalyzed reactions that constitute a biochemical pathway. Then, in a second step, one might want to study interactions between organelles or cells, and finally, the focus could be the dynamics of a system within its environment.

This hierarchy of foci raises the question, To what degree can the structure and results of lower-level models be incorporated into higher-level models? In other words, if a single variable of the higher-level system constitutes an entire system at the lower level, how do the two models relate to each other? In linear models, where the right-hand sides of all differential equations consist of sums of variables, this substitution of a variable with a lower-level system does not change the linear structure of the higher-level model. However, as soon as the differential equations contain nonlinear terms, the structure is usually destroyed. A very rare exception is found in S-systems and related differential equations based on products of power functions. Like linear models, S-systems preserve their structure when variables are exchanged for lower-level systems (58). Because one can focus on different levels of the same phenomenon and always use S-system models of the same type, one says that S-systems have telescopic properties.

**Theoretical justification.** No model can ever be proven mathematically to be correct. A model is based on abstractions and simplifications that on one hand make the model easier to understand than the modeled reality, but on the other hand result in differences between model responses and reality. Depending on the type of abstractions and simplifications, the compromise between validity and simplification turns out differently and it is difficult, if not impossible, to rank competing models according to their overall quality.

In addition to its consistency with experimental observations, the S-system form is supported by different types of theoretical arguments.

- S-systems are derived from a very general system description through rigorous mathematical methods of approximation theory. This theory assures us that S-systems are valid representations if the concentrations of all variables remain relatively close to their nominal operating values. Very often this is indeed the case. Many *in situ* systems are very well buffered against variations in concentration. Experience with a variety of biologic and nonbiologic phenomena has further shown that the products of power functions, on which the S-system equations are based, are in fact valid representations. Often, a metabolite concentration can be varied 10 times or 100 times and still the power-function representation is sufficiently accurate. These ranges of variation are much wider than typically seen *in situ*. The most prominent and ubiquitous mechanism that tends to keep concentrations within close limits is feedback inhibition, but other mechanisms like feedforward regulation and the physiologic shortening of pathways [for references, see (7)] are also powerful means for buffering concentration variations *in vivo*. These observations and theoretical explanations support the use of S-system representations in environmental analyses.
- Every mathematical model can describe a certain repertoire of behaviors, whereas other behaviors cannot be modeled. For instance, monotonic functions cannot represent oscillations, and linear differential equations cannot describe saturation. It is often not easy to determine which types of behaviors particular nonlinear differential equations can capture and what is outside their reach. In the case of S-systems, it was shown with mathematical rigor that virtually any phenomenon that can be formulated as a differential equation at all can also be formulated as an equivalent S-system. For instance, oscillations of numerous types can be modeled, including limit cycle oscillations that cannot be represented with linear systems. Even deterministically chaotic systems such as Lorenz and Rössler oscillators are special cases of canonical models. This richness of canonical models was demonstrated by a constructive proof showing how differentiable functions and ordinary differential equations can be recast equivalently in the form of canonical S-system, GMA system, or half-system models (29). This demonstration is of importance, as it assures the canonical modeler that there are no structural limitations to what these models can represent. By contrast, linear models provide the great advantage of



uncounted algebraic methods of analysis, but because these models cannot capture important phenomena like saturation or stable oscillations, linear models are of limited use by their very nature.

**Analytic convenience.** S-system models have unique properties that make mathematical and numerical analyses possible and often relatively simple.

- For S-system models, the translation of a flow diagram of a real-world phenomenon into differential equations follows straightforward recipes. This is possible because the differential equations always have the same homogeneous structure and because all parameters have a well-defined meaning. In contrast, *ad hoc* models require a high degree of mathematical ingenuity and often numerous assumptions and simplifications that may be hard to justify. For instance, there are no generally applicable rules for devising the optimal mechanistic rate law for a highly regulated pathway *in vivo*. Furthermore, the complexity of *ad hoc* models grows immensely with the inclusion of additional metabolites and regulators, whereas the structure of the S-system differential equations always remains the same. Of course, the number of variables, equations, and parameters grows, but even for large systems it is still possible to set up the equations with the same simple recipes.
- All parameters of an S-system model have clearly defined meanings. This is very important. In principle, all parameter values can be deduced from the environmental system through experimental measurements. These experiments may be difficult to execute, or one may not yet have the techniques to perform them at all. However, in contrast to "fudge factors," which are mathematically necessary but have no actual meaning and therefore never will be measurable, time and scientific ingenuity will provide the technology necessary to estimate S-system parameters.
- The particular mathematical form of the S-system differential equations offers outstanding features when it comes to mathematical and computer-aided analysis. No other general type of differential equations is known to provide realistic and relevant representations of the same mathematical convenience.
- The homogeneous structure of S-systems permits comparative analyses of alternate model structures that are conceptually analogous with control experiments in the bench sciences. These mathematically controlled analyses elucidate the relative importance of any of the system components and explain why nature has selected observed regulatory patterns rather than

other possible designs (59–64). Even though this type of mathematical experimentation provides more insight than many other typical analyses, it is not often performed in other approaches to systems analysis, because mathematical models without a homogeneous structure are not well suited for such mathematically controlled experiments.

## Dynamics

Once a canonical model is set up and its parameters are determined, two types of analyses can be performed. One explores features of the dynamics of the system, that is, of changes in the values of variables over time. The other type of analysis addresses features associated with the steady state of the system, in which the variables do not change in magnitude and all influxes and effluxes are in balance. Steady-state analyses are discussed in the following section.

Typical dynamic experiments with canonical models fall in to three categories. First, one begins with a system at steady state, perturbs one of the dependent variables, and studies the responses of the system over time. For instance, one may study whether the system returns to the original steady state, and whether the transient behavior is monotone or shows overshoots, undershoots, or oscillations. In a population model, this type of experiment could correspond to the accidental death of many individuals. In a biochemical model, it could correspond to the exogenous supply of one of the metabolites of interest.

The second type of experiment changes the value of one of the independent variables. Such a change is truly different from the first type of experiment, since the system has no means of adjusting the independent variable. Thus, the alteration is permanent and not transient as it is when a dependent variable is changed. If the system reaches a steady state after the perturbation, it is usually a state different from that where the system started.

The third type of experiment deals with changes in parameter values. Again, these changes are permanent, as the system does not have the capability of changing back the parameter values.

**Special cases.** Instead of demonstrating these types of experiments with generic examples, it might be more useful to review special cases in the context of environmental health. The interested reader may consult the literature for detailed numerical examples of dynamic analyses of real-world systems (65–68).

**Physiologically based pharmacokinetic models.** Physiologically based pharmacokinetic models (PBPK) belong to the standard repertoire of risk assessment (69). These models are designed to investigate how chemicals

distribute among the compartments within an organism over time. For instance, upon an intravenous injection of a chemical, this chemical is distributed throughout the bloodstream and enters the various organs at rates determined by their volumes, flow rates, and other physiologic features. The models capture the temporal changes in concentrations in each compartment.

PBPK models typically use differential equations with first-order kinetics for transport and Michaelis-Menten functions for enzyme-catalyzed reactions within compartments. Although these functions are often sufficient, they are not always valid *in situ* (8). As a specific example, Krishnan et al. (70) found in a PBPK analysis of chemical mixtures that simple first-order or Michaelis-Menten kinetics did not always capture the observed dynamics of dichloroethylene in rats, and hypothesized that some of the involved enzymes were subject to inhibition. Indeed, the inclusion of inhibitory effects improved the data fit considerably. Accounting for inhibition in traditional enzyme kinetics requires knowledge or assumptions about the mechanism of inhibition, and the chosen mechanism dictates the mathematical format of the equation. For simple systems, this choice may not be difficult, but once PBPK models approach the level of sophistication of some of the more advanced models of biochemical systems, the use of generalized Michaelis-Menten models becomes limiting [for example, see models and comparative discussions in (71–73)]. An alternative is the use of canonical models, in which inhibitions and other regulatory signals are formulated in the streamlined fashion of power-law functions. The resulting representations of metabolic fluxes or clearance terms are readily implemented as stand-alone models or within the shell of a PBPK model (74).

**Multistage models.** These popular models of carcinogenesis have not been approached with methods of canonical model analysis, but they do have the form of simple GMA systems (75,76). Typically, multistage models are analyzed only for very low doses of radiation or chemical insults, which allows simplifications up to a point where the probability of a resulting malignancy is a linear function of dose. Nevertheless, the full multistage model is dynamic and methods of canonical modeling could be applied for its analysis.

**Ecosystem models.** These models are typically formulated as Lotka-Volterra models (30,31,33). They consist of a system of ordinary differential equations in which the right-hand side of any variable  $X_i$  consists of a product of  $X_i$  itself with a linear combination of all dependent variables, i.e.:

$$\dot{X}_i = X_i \cdot \sum_{j=1}^n (a_{i0} + a_{ij}X_j). \quad [22]$$



Lotka-Volterra systems are direct special cases of GMA systems in which each product consists of one or two factors with exponents of 1.

**Transient mass balance models.** These models are frequently used to describe the flow of material through ecosystems. They often take the general form

$$\frac{d(\text{stored mass})}{dt} = \text{transport in} + \text{source production} - \text{transport out} - \text{sink elimination} \quad [23]$$

(77). Clearly, these systems fall into the general system description (Equation [3]), and can thus be approximated by canonical models, as shown above. On a much smaller scale, the same type of model may describe the change in chemical mass with respect to time at any small unit of the two- or three-dimensional space.

**Survival.** For simplicity of argument, suppose one is interested in a cohort of individuals without replenishment. The survival dynamics of this group is described by a curve that starts at 100% and monotonically decreases toward 0. Survival phenomena are very complex because they are not only functions of time but also depend on uncounted internal and external factors and processes that ultimately lead to survival or death. In light of this complexity, it is amazing that the overall appearance of observed survival curves is usually rather smooth and simple.

A possible explanation for this phenomenon can be found in a mathematical deduction for the relative simplicity of growth curves, as proposed by Savageau (58). Suppose in a *Gedankenexperiment* that all processes contributing to survival could be formulated in a comprehensive, multi-variate S-system model. Since S-systems can have essentially any degree of complexity, such a model would theoretically be an option, even though one would not be able in practice to implement it.

Given the complex nature of survival, the involved processes are manifold, and it is likely that they run at vastly different time scales. At one extreme, biochemical and molecular processes occur within seconds or minutes, while at the other extreme, evolutionary processes and global climate changes are very slow in comparison to the life span of an individual or population. As discussed above, variables describing very fast or very slow processes can be replaced with constants. The very fast processes reach their steady states so quickly that they are essentially always in steady state, whereas the very slow processes change so little throughout the lifetime of the individual or population that the change is

negligible. In either case, the time derivative of each associated variable is 0, and its differential equation becomes an algebraic function, constraining the remaining variables. This function can again be approximated as a power-law function and substituted in the S-system, which retains its mathematical structure but is reduced in size.

Thus, the survival process is governed by only a few processes that occur at just the right time scale. If survival processes are dominated by just one or two dominating hazards, they can ultimately be represented with one- or two-variable S-systems. Savageau (54,58,78) supported the analogous conclusion for growth functions by demonstrating that many published growth laws are in fact one- or two-variable S-systems. The same type of support is available for statistical distributions and survival functions; many of them are well represented by small S-systems or even a single S-system equation (79–82). This so-called S-distribution is expressed in terms of the random variable  $X$  and its cumulative distribution  $F$  and takes the form.

$$\frac{dF}{dX} = \alpha(F^g - F^h) \quad F(X_0) = 0.5. \quad [24]$$

$X_0$  is the median of the distribution, the positive parameter  $\alpha$  determines its spread, and the real-valued powers  $g$  and  $h$  ( $g < h$ ) characterize its shape. In survival analysis, the random variable  $X$  represents time ( $X = t$ ) and  $F$  represents the cumulative failure distribution, which is the complement of the survival function  $S(t)$ :  $F(t) = 1 - S(t)$ .

S-distributions have interesting properties from an academic as well as a practical point of view. They provide a shape-based classification of traditional and new distributions (80) and a convenient tool for Monte-Carlo simulations in which input parameters have distributions whose mathematical structure is a priori unknown (83,84). They also offer a number of procedures for parameter estimation and random number generation, which in turn, is an essential component of any Monte-Carlo simulations.

### Steady-State Analysis

**Computation of steady states of S-systems.** Many aspects of a dynamic model are related to the steady state, in which production and depletion for all variables are in balance. In contrast to dynamic analyses that address the temporal features of system responses to perturbations, steady-state analyses deal with more permanent changes in the system characteristics. A typical example is a persistent change in environmental conditions that evokes a permanent change in the values of some or all the variables.

The steady state of any dynamic model is simply characterized by the equations

$$\dot{X}_i = 0 \quad i = 1, 2, \dots, n. \quad [25]$$

For nonlinear systems, these equations cannot usually be solved with algebraic means and require solution by some search algorithm. A notable exception is the model formulation as an S-system. In this representation the steady-state equations actually become linear, and this facilitates a host of further analyses. The linearity of the steady-state equations of S-systems is readily demonstrated. Upon setting the differential equations equal to 0, the  $\beta$ -terms are brought to the left-hand side. Taking logarithms on both sides and sorting terms yields a set of linear equations in the logarithms of the original dependent and independent variables.

Thus, for a system without independent variables, one obtains,

$$0 = \alpha_i \prod_{j=1}^n X_j^{g_{ij}} - \beta_i \prod_{j=1}^n X_j^{h_{ij}} \quad i = 1, 2, \dots, n, \quad [26]$$

or, equivalently,

$$\alpha_i \prod_{j=1}^n X_j^{g_{ij}} = \beta_i \prod_{j=1}^n X_j^{h_{ij}} \quad i = 1, 2, \dots, n. \quad [27]$$

Given the most relevant case that none of the rate constants  $\alpha_i$  and  $\beta_i$  are zero, one can take the logarithm on both sides and obtain

$$\ln(\alpha_i) + \ln\left(\prod_{j=1}^n X_j^{g_{ij}}\right) = \ln(\beta_i) + \ln\left(\prod_{j=1}^n X_j^{h_{ij}}\right) \quad [28]$$

Because the logarithm of a product converts into a sum of logarithmic terms, the steady-state equations become

$$\begin{aligned} \ln(\alpha_i) + \sum_{j=1}^n [g_{ij} \ln(X_j)] \\ = \ln(\beta_i) + \sum_{j=1}^n [h_{ij} \ln(X_j)] \quad i = 1, 2, \dots, n. \end{aligned} \quad [29]$$

Now define  $y_i = \ln(X_i)$  and move all terms containing  $y_i$ s to the left side and all terms without  $y_i$  to the right side. The result is

$$\sum_{j=1}^n g_{ij} y_j - \sum_{j=1}^n h_{ij} y_j = \ln(\beta_i) - \ln(\alpha_i) \quad i = 1, 2, \dots, n. \quad [30]$$

The only step left is to rename  $a_{ij} = g_{ij} - h_{ij}$  and  $b_i = \ln(\beta_i) - \ln(\alpha_i) = \ln(\beta_i/\alpha_i)$  for all  $i$  and all  $j$ . With that, a general S-system with  $n$  dependent variables and no independent

variables has a steady state characterized by a set of  $n$  linear equations of the form

$$\begin{aligned} a_{11}y_1 + a_{12}y_2 + a_{13}y_3 + \dots + a_{1n}y_n &= b_1 \\ a_{21}y_1 + a_{22}y_2 + a_{23}y_3 + \dots + a_{2n}y_n &= b_2 \\ a_{31}y_1 + a_{32}y_2 + a_{33}y_3 + \dots + a_{3n}y_n &= b_3 \\ &\vdots \\ a_{n1}y_1 + a_{n2}y_2 + a_{n3}y_3 + \dots + a_{nn}y_n &= b_n. \end{aligned} \quad [31]$$

and this is true as long as no rate constant is 0. In matrix notation, Equation 31 is written succinctly as

$$A \cdot \vec{y} = \vec{b}. \quad [32]$$

Matrix  $A$  consists of the coefficients  $a_{ij} = g_{ij} - h_{ij}$ ,  $\vec{y}$  is a row vector consisting of the  $n$  components  $y_j = \ln(X_j)$  ( $j = 1, \dots, n$ ), and  $\vec{b}$  is a row vector containing the solution coefficients  $b_i = \ln(\beta_i/\alpha_i)$  ( $i = 1, \dots, n$ ).

When there are  $m$  independent variables  $X_{n+1}, \dots, X_{n+m}$  exactly the same procedures apply. As before, one introduces new variable names  $y_j = \ln(X_j)$ , for  $j = 1, 2, \dots, n+m$ , and defines coefficients  $a_{ij} = g_{ij} - h_{ij}$  and  $b_i = \ln(\beta_i/\alpha_i)$ . The result consists of  $n$  linear equations in  $n+m$  variables (cf. Equation 31):

$$\begin{aligned} a_{11}y_1 + a_{12}y_2 + \dots + a_{1n}y_n \\ &= b_1 - a_{1,n+1}y_{n+1} - \dots - a_{1,n+m}y_{n+m} \\ a_{21}y_1 + a_{22}y_2 + \dots + a_{2n}y_n \\ &= b_2 - a_{2,n+1}y_{n+1} - \dots - a_{2,n+m}y_{n+m} \\ a_{31}y_1 + a_{32}y_2 + \dots + a_{3n}y_n \\ &= b_3 - a_{3,n+1}y_{n+1} - \dots - a_{3,n+m}y_{n+m} \\ &\vdots \\ a_{n1}y_1 + a_{n2}y_2 + \dots + a_{nn}y_n \\ &= b_n - a_{n,n+1}y_{n+1} - \dots - a_{n,n+m}y_{n+m} \end{aligned} \quad [33]$$

Note that the equations are arranged in such a way that the (unknown) dependent variables are separated from the (known) independent variables. The left-hand side is exactly the same as the expression  $A \cdot \vec{y}$  that was developed above for systems without independent variables. The right-hand side contains the solution vector  $\vec{b}$  as before but also the independent variables and coefficients associated with independent variables. Distinguishing vectors and matrices of coefficients associated with dependent or independent variables with indices  $D$  and  $I$ , one obtains the steady-state matrix equation

$$A_D \cdot \vec{y}_D = \vec{b} - A_I \cdot \vec{y}_I. \quad [34]$$

If a steady-state solution exists in which no variables are 0, the matrix  $A_D$  has maximal

rank, and one can use the inverse  $A_D^{-1}$  to express the solution formally as

$$\vec{y}_D = A_D^{-1} \cdot \vec{b} - A_D^{-1} \cdot A_I \cdot \vec{y}_I. \quad [35]$$

This formulation demonstrates explicitly how the dependent variables are affected by the system characteristics, as represented by the coefficients  $a_{ij} = g_{ij} - h_{ij}$  in the matrices  $A_D$  and  $A_I$ , the solution coefficients  $b_i = \ln(\beta_i/\alpha_i)$ , and the independent variables, given in logarithmic coordinates as  $\vec{y}_I$ .

GMA systems can also exhibit steady states, which are characterized by the equations

$$\dot{X}_i = 0 \quad [36]$$

for all dependent variables. However, since the differential equations of GMA systems do not necessarily have exactly one difference of products of power-law functions on the right-hand sides, their steady states cannot be converted into linear equations as in S-systems. Steady-state solutions of GMA systems therefore require numerical search algorithms.

**Stability and sensitivities.** The explicit representation of steady states of S-systems of arbitrary size is a great advantage for further analyses. Among the standard diagnostics for steady states are assessments of stability and sensitivities. The concept of stability can be divided further into two classes: local and structural. Local stability assesses whether a system will return to the original steady state after a small perturbation, whereas structural stability deals with the question of whether the qualitative behavior of the system changes if one of the parameters is altered. For instance, it may happen that a system begins to oscillate if a parameter value is increased or decreased.

For S-system models, questions of local stability can be answered algebraically or numerically with well-known methods of eigenvalue analysis (7,23,27,85). Structural issues are generally much more complicated, but for the widely relevant case of the emergence of limit cycle oscillations (at Hopf bifurcations), the S-system structure provides astonishingly simple criteria (86).

Sensitivity analysis also addresses questions of structure and robustness. The key idea here is to quantify the magnitude of a system response to small changes in a parameter value or independent variable. If the system responds strongly to minute changes, it is usually deemed unrealistic or unreliable, as small variations in parameters or independent variables are encountered in the real world on a regular basis. Because the steady-state equations of S-systems are linear, methods of linear algebra are directly applicable to the analysis of sensitivities. Even though GMA systems and other nonlinear models do not

have linear steady-state equations, sensitivities can be computed, for instance, with methods of implicit differentiation (87,88).

**Special cases of steady-state analyses in environmental health assessment.** As in the previous section, it might be useful to study some steady-state features of canonical models in the context of environmental health by a review of special applications.

**Exposure models.** Because of their sizes and multitudes of details, exposure models often appear to be rather complex. However, in the majority of cases, they are ultimately sums of products. The products quantify parameters like dose, duration, body weights, and such aspects as bioavailability, while the sum of these products accounts for different exposure routes and different exposure scenarios. More complicated models of exposure assessment account for spatial and temporal aspects of source and contact.

A typical example, in rather general terms, is an exposure equation that accounts for differences in concentrations and exposure durations among different microenvironments (89),

$$E_{total} = \sum_i f_i (P_i C_{amb} + S_i). \quad [37]$$

Each partial exposure in a microenvironment is determined by the fraction  $f_i$  of time spent in the  $i$ th microenvironment, the ambient concentration  $C_{amb}$ , the effective penetration factor  $P_i$  for the ambient pollution into the microenvironment, and the effective source strength  $S_i$  of the pollutant.

Suppose one constructs an exposure model within the framework of canonical GMA systems. The variables would include the concentration of the target chemical (the agent), which one could code as  $X_1$ , as well as all kinds of physical features at the location of exposure, which could be coded as  $X_2, \dots, X_n$ . A dynamic simulation with the model would describe how the agent and some of the physical features change over time. For simplicity of argument, one may assume that  $X_1$  decays according to a first-order process and that it has no other direct effect on its own dynamics. This assumption may or may not always be justified but is in line with traditional approaches to exposure assessment. The GMA equation of  $X_1$  in this situation reads

$$\begin{aligned} \dot{X}_1 &= \gamma_{11} \prod_{j=2}^n X_j^{f_{11j}} \pm \gamma_{12} \prod_{j=2}^n X_j^{f_{12j}} \pm \dots \\ &\pm \gamma_{1k} \prod_{j=2}^n X_j^{f_{1kj}} - \gamma X_1, \end{aligned} \quad [38]$$

where none of the products contain  $X_1$ .

At steady state, Equation 38 is set equal to 0 and divided by the positive rate  $\gamma$ . Bringing  $X_1$  to the left-hand side results in a

steady-state exposure equation that consists of a sum of products of power-law functions. The same is true if the degradation term  $\gamma X_1$  is replaced with a full product of power-law functions. In this case, the steady-state equation is first divided by this product, except for  $X_1$ .

If all variables affect exposure in a linear fashion, all exponents  $f_{ij}$  are either 0 or 1, and the steady-state exposure equation reduces to the traditional form. Thus, simplifying the full GMA model in such a way that it reflects traditional assumptions produces the well-known results, quasi as a special case. The full GMA model goes beyond the traditional model by accounting for nonlinear effects, synergisms, and feedbacks, and of course the overall dynamics of the exposure process.

**Responses to low-dose chemical exposure.** For this scenario, consider how an environmental agent affects the metabolism of an exposed organism. It may influence one or several pathways and act in different ways; for instance, as substrate for one pathway and inhibitor of another. Consistent with the philosophy of canonical modeling, the agent is an independent variable, as its concentration outside the body and its dynamics in the environment are not directly affected by the metabolic activity within the organism.

Suppose one could model all pathways directly or indirectly affected by the environmental agent and formulate this model as a potentially huge S-system. The dependent variables of this system would be metabolites, and the agent of interest would be among the independent variables. Whereas some clinical symptom presumably would be the ultimate measure of health risk, it is useful to study biochemical responses, which eventually could lead to an adverse health outcome. At the biochemical level, the typical response to the presence of the agent is a combination of sustained, elevated, or decreased concentrations or fluxes. In terms of canonical modeling, the system (temporarily) assumes a new steady state, and a comparison of this state with the original state is a measure for the potency of the agent.

To be specific, suppose the system consists of  $n$  dependent variables, the agent of interest is represented by the independent variable  $X_{n+k}$ , and the metabolic response of most interest occurs in the dependent variable  $X_j$ . If  $X_{n+k}$  is elevated, the system reacts by assuming a new steady state in which the response or (response rate) is increased (or decreased), leading to an elevated (or lowered) steady-state value of  $X_j$ . Of course,  $X_{n+k}$  also affects other variables, and because of the connectivity of the organism's metabolism,  $X_j$  is affected secondarily by these changes in other variables.

Because there are independent variables, the new steady state of the S-system model is characterized by a set of linear equations with more variables than equations, as shown above. Limiting the analysis exclusively to the effect of the environmental agent, coded in logarithmic form as  $y_{n+k}$ , on the dependent variable of interest, coded in logarithmic form as  $y_j$ , one thus obtains the result

$$y_j = c_1 + c_2 y_{n+k} \quad [39]$$

where  $c_1$  and  $c_2$  are well-defined linear combinations of many of the original system parameters. The parameter  $c_2$  is known in the literature as logarithmic gain (7,23,27,90).

The linear relationship is expressed in logarithmic coordinates but readily translated back to the Cartesian space. The result asserts that the dependent variable of interest,  $X_j$ , is a power-law function of the independent variable  $X_{n+k}$ :

$$X_j = \exp(c_1) \cdot X_{n+k}^{c_2} \quad [40]$$

This result includes not only the direct effect of the agent on  $X_j$  but all indirect effects that filter through the potentially huge, nonlinear system.

Thus, whatever singular or multiple effects an environmental agent exerts upon a metabolic system, each dependent variable responds in a power-law fashion as long as the perturbation is not too large. The responses generally differ among the dependent variables in magnitude and direction, with some increasing and some decreasing, but the mathematical structure of their responses is always the same, namely that of a power-law function. Further details are presented elsewhere (91).

**The linear-logistic model.** Epidemiologists commonly approach health risks to populations by studying the proportion,  $P$ , of the diseased individuals,  $D$ , within a population of size  $N$ :  $P = D/N$ . More specifically, one uses the concept of the odds of this proportion. For the relevant case where not all individuals in the population are healthy, this odds is defined in terms of diseased and healthy ( $H$ ) individuals as

$$\frac{P}{1-P} = \frac{(N-H)/N}{1-(N-H)/N} = \frac{D}{H} \quad [41]$$

[cf. (92)]. The linear-logistic model assumes that the proportion  $P$  is related to a weighted sum of risk factors  $Y_i$  and takes the form of a multi-variate logistic function:

$$P = \frac{1}{1 + e^{-(a_0 + a_1 Y_1 + a_2 Y_2 + \dots + a_n Y_n)}} \quad [42]$$

The coefficients  $a_1, \dots, a_n$  quantify the impact of the respective risk factors on disease development, whereas the coefficient  $a_0$  is not associated with a particular risk factor but reflects the background risk. It corresponds to the logarithm of the odds for disease among the unexposed.

In light of the enormous complexity that governs the dynamics of a disease in a population, it is quite surprising that the rather simple linear-logistic model turns out to be widely applicable and accurate. The canonical approach can provide some explanations why this may be so.

Suppose the disease process is formulated as a (potentially very large) S-system model. For bookkeeping purposes, the two variables  $X_{n+1}$  and  $X_{n+2}$  of the S-system code for the numbers of diseased and healthy individuals, respectively. At steady state, the system can be solved and leads to a set of linear equations in logarithmic coordinates, as shown above. Translation back to Cartesian coordinates produces the corresponding solution in terms of products of power-law functions. Specifically, one obtains

$$\begin{aligned} D = X_{n+1} &= \gamma_{n+1} X_1^{e_1} X_2^{e_2} \dots X_n^{e_n}, \\ H = X_{n+2} &= \gamma_{n+2} X_1^{f_1} X_2^{f_2} \dots X_n^{f_n}, \end{aligned} \quad [43]$$

where the subscripted parameters  $\gamma$ ,  $e$ , and  $f$  are sums and products of the original S-system parameters (21,24). Substitution of  $D$  and  $H$  in to Equation 41, the formula for the odds of the disease proportion, results in

$$\frac{P}{1-P} = \frac{\gamma_{n+1}}{\gamma_{n+2}} X_1^{e_1-f_1} X_2^{e_2-f_2} \dots X_n^{e_n-f_n} \quad [44]$$

Simple renaming  $a_0 = \ln(\gamma_{n+1}/\gamma_{n+2})$ ,  $a_j = e_j - f_j$ ,  $Y_j = \ln(X_j)$  for  $j=1, \dots, n$  yields

$$\frac{P}{1-P} = \exp(a_0) \exp(a_1 Y_1) \exp(a_2 Y_2) \dots \exp(a_n Y_n) \quad [45]$$

Algebraically solving for  $P$  readily shows that Equation 45 is equivalent to the linear-logistic model in Equation 42. This result is very interesting because it implies that the linear-logistic model is a natural and necessary consequence of the formulation of the disease process as a dynamic model in S-system form. The result derives from the canonical model essentially without assumptions beyond general tenets of dynamic modeling and the power-law approximation that underlies the canonical approach. It is not only academically intriguing but also has some practical implications. On one hand, one can interpret the parameters in the linear-logistic model as weights of the contributing risk factors, as

done in epidemiology. On the other hand, it is now also possible to interpret them within the framework of dynamic modeling, where the weights correspond to amplification or gain factors (7,24,90) that are very well known in engineering and general systems theory.

All standard computations with the linear-logistic model can be executed within the framework of canonical models (24). For instance, one can produce  $2 \times 2$  tables and compute the proportions of diseased individuals exposed to some environmental agent or who possess a given risk factor. Of course, the results are identical, but the canonical modeling framework allows additional interpretations and an assessment of the limitations of the linear-logistic model.

The derivation of the linear-logistic model from concepts of systems theory provides a natural explanation of the multiplicativity of risk factors: In the underlying S-system model, the risk factors appear as products, and the strengths of their effects are given as powers. If such a power is positive, the effect is positive, if it is negative, the effect is negative, and a power of zero signifies the irrelevance of the factor in question. If a risk factor affects only one process within the system, the associated power in the S-system model translates directly into the corresponding coefficient of the linear-logistic risk model. However, if a risk factor affects several processes, the corresponding coefficient in the linear-logistic model combines all effects in a fashion that is not *a priori* evident but that directly follows the simple algebra presented above. The coefficient of the linear-logistic model thus loses some resolution over the dynamic model. It reflects the overall effect of the risk source but, for instance, cannot distinguish between direct and indirect effects or between modulations of the birth rate or the death rate.

One may ask what happens if the general disease model is not formulated as an S-system model. The overall answer is that the odds of the proportion are not likely to be in the form of the traditional linear-logistic model. In fact, most nonlinear models do not permit the formulation of explicit steady-state equations, and this makes further evaluations difficult if not impossible. A definite answer can be given if the general disease model is approximated by a linear model, which itself is canonical in nature. In this case, the dynamic equations for the diseased and healthy subpopulations read

$$\begin{aligned} D &= X_{n+1} = a_{n+1,0} + a_{n+1,1}X_1 + a_{n+1,2}X_2 \\ &\quad + \dots + a_{n+1,n+2}X_{n+2}, \\ H &= X_{n+2} = a_{n+2,0} + a_{n+2,1}X_1 + a_{n+2,2}X_2 \\ &\quad + \dots + a_{n+2,n+2}X_{n+2}. \end{aligned} \quad [46]$$

The associated steady-state equations are linear, and barring unusual singularities, one can express  $X_{n+1}$  and  $X_{n+2}$  as linear functions of  $X_1, \dots, X_n$ . The proportion of diseased,  $P = D/N$ , also is a linear function, since  $N$  is a constant at steady state. Thus, the linear disease model leads to the linear risk model

$$P = \alpha + \sum_{i=1}^n \beta_i X_i. \quad [47]$$

This type of risk model indeed is sometimes found in textbooks. It is to be used as an alternative to the linear-logistic model in cases in which the risk difference is of interest (93).

From a systems point of view, there is exactly one decision that needs to be changed to obtain the linear model instead of the linear-logistic model. This decision is to approximate the general system model (Equations 1–3) not with an S-system model (i.e., via linearization in logarithmic coordinates) but with a linear model (i.e., via linearization in Cartesian coordinates), which emphasizes the additive rather than the associative nature of the processes that govern the system.

**Cox's proportional hazards model.** The use of the linear-logistic model becomes problematic if significant numbers of individuals are lost to follow-up or to unrelated death. It is instead customary to base risk estimates on person-years of survival and to compute an incidence rate, incidence density, or hazard rate, as originally suggested by Cox (94). The resulting model typically takes the form

$$\begin{aligned} \ln(\text{incidence rate}) \\ = b(t) + b_1 Y_1 + b_2 Y_2 + \dots + b_n Y_n. \end{aligned} \quad [48]$$

The term  $b(t)$  describes a time-dependent baseline effect, whereas the coefficients  $b_1, \dots, b_n$  account for the contribution of each risk factor to the overall disease incidence rate. As in the case of the linear-logistic model, this relationship is rather simple, yet surprisingly accurate and applicable, and again, the corresponding canonical model provides some explanations why that is so. The dynamics of the group of diseased individuals is given by the equation

$$\dot{X}_{n+1} = \alpha_{n+1} \prod_{j=1}^{n+2} X_j^{g_{n+1,j}} - \beta_{n+1} \prod_{j=1}^{n+2} X_j^{h_{n+1,j}}. \quad [49]$$

Furthermore, the instantaneous increase in the number of diseased individuals is specifically and uniquely represented by the  $\alpha$  term of this equation. Thus, this term is equivalent to the incidence rate in question. Taking the logarithm of the  $\alpha$  term and recoding  $Y_i = \ln(X_i)$  as before, one directly obtains

$$\ln(\text{incidence rate}) \propto \ln(\alpha_{n+1}) + \sum g_{n+1,j} Y_j \quad [50]$$

which already has almost the form of Cox's model in Equation 48. The only apparent difference is the time dependence of the baseline term in Cox's model,  $b(t)$ . This difference is readily explained. It is very unlikely that absolutely all risk factors are captured in Cox's risk model. Thus, there are variables that actually contribute to the disease process but are not explicitly listed. They may represent physiologic or environmental factors that drive the population dynamics, regardless of whether the population is exposed to the risk factors in question. These variables are time dependent, but as they are not explicitly stated in Cox's model, they are lumped with  $\alpha_{n+1}$  in the corresponding canonical model, thereby making this parameter a time-dependent term that corresponds to  $b(t)$ . As in the case of the linear-logistic model, these results follow directly from the S-system description of the disease process.

**Importance of subsuming special cases.** A crucial feature of any theoretic framework is the assessment of its boundaries. In pure mathematics, boundaries and assumptions are explicitly stated with every definition and every theorem. In an environmental context, a precise definition of boundaries is often difficult to obtain and therefore substituted with collective positive experience in similar situations. In the present context, the linear-logistic model and Cox's proportional hazard model have been used successfully many times, even though it was not at all clear why these models were appropriate and what their boundaries were.

The demonstration that these and other environmental models are special cases of canonical S-system models provides a theoretic framework that permits explanations and further assessments. All assumptions that explicitly or implicitly underlie the environmental models can now be traced back to the assumptions that underlie the canonical approach and its implementation in the form of S-systems. This strong result is true, as no assumptions beyond those underlying the canonical approach are necessary to derive these environmental models. In particular, questions of appropriateness and validity of the models are mapped onto questions of approximation quality. As long as the power-law approximation of the disease model is valid, the epidemiologic models are valid.

The weights in the risk functions of the previous section can be reinterpreted as logarithmic gains, which are sensitivities of the model responses to changes in independent variables. There exists a large body of literature on sensitivity analysis, and because of the re-interpretation of weights as sensitivities, this knowledge can now be brought to bear on the epidemiological models.

The sensitivities indicate how a dependent variable responds to changes in independent variables. In an environmental health context, they describe the instantaneous change in the number of sick that is a consequence of a change in a risk factor. Sensitivities are conceptually based on infinitesimally small variations in independent variables. This implies that the coefficients in the two epidemiologic models also are, strictly speaking, guaranteed only for infinitesimally small changes in risk factors. Experience with sensitivities of canonical models has shown that responses to variations of 5%, 10%, or more in an independent variable are usually quite accurately captured by sensitivity analysis. However, there is no guarantee, and in a particular application one may significantly over- or underestimate the responses. Because of the equivalence between risk-factor weights and sensitivities, one has to expect analogous results for the epidemiologic models. For instance, the linear-logistic model breaks down for large extrapolations in risk factors and in situations in which risk factors are not independent. Assessing these observations from a canonical modeling point of view, one finds new interpretations of these failures of the epidemiologic models and can develop modeling strategies for avoiding them.

For instance, if dependencies are included in a linear-logistic model, one typically represents them in the risk equation as multiplicative terms like  $X_2X_3$ . The S-system model, by virtue of its structure, directly accounts for interactions between variables, and the resulting risk equation does not directly allow for such terms. However, if one assumes that one of the kinetic orders is a function of time or of other variables, the above derivation leads directly to multiplicative terms. As an alternative, one may define an additional variable that represents a product like  $X_2X_3$ . The theory behind S-systems then guarantees that one can again write the differential equation of this additional variable as an S-system equation. Thus, according to the above derivation, the additional variable, and thus the product  $X_2X_3$ , appears in the risk equation like any original variable.

## Conclusion

Mathematical modeling is a relatively young field. It has not matured to a point where there is a canon of good professional practices guiding the modeler from a simple verbal or graphic description of the phenomenon to the design, analysis, and interpretation of a generally accepted model. Some methods, such as assessments of stability and robustness are becoming standard procedure, but beyond these basic features, the choices of model structure and features to be analyzed are rather wide.

Canonical modeling provides guidelines for setting up models and for analyzing and interpreting them. At first, canonical models seem to be overly restricting straightjackets, as they allow only one type of function, namely sums and differences of products of power-laws. However, it has been shown that these types of representations are rich enough to model all dynamic behaviors the biologist or environmental health expert is likely to encounter (29). Furthermore, a wide array of applications has demonstrated the usefulness and power of these types of models.

This review has attempted to demonstrate how canonical modeling may play a role in integrated assessments of environmental health issues at various levels. Canonical models of biochemical and metabolic systems have the potential of explaining, at the level of metabolites and enzymes, how toxic agents affect an organism and why different species may respond differently to comparable exposures. Similar models may be good candidates for quantifying and explaining the dynamics of biomarkers, whether they are used as diagnostic tools of exposure or in toxicity testing. Numerous canonical analyses have elucidated patterns of gene expression and regulation (62–64). At a somewhat higher level of organization, canonical models have been used at the boundary of physiology and population dynamics (83,95,96). At the population level itself, canonical analyses have addressed the boundaries between dynamic modeling and statistics. The assessment of disease patterns with the linear-logistic and Cox models was shown as an example above. Other studies demonstrated how dynamic models can lead to trends in statistical distributions (97). Examples in this category include the metabolic accumulation of contaminants in fish (83) and the assessment of growth patterns in children (98). Finally, canonical models were used to describe transitions from dynamic models to survival processes (26).

At this point in time, canonical model analyses have addressed systems with moderate numbers of components. For instance, Curto et al. (68) constructed models of purine metabolism with 18 variables, Shiraishi and Savageau (72) designed a model of the TCA cycle consisting of about 50 variables, and Ni and Savageau (67,73) proposed a model for red blood cell metabolism involving close to 100 variables. These sizes of models are an order of magnitude more complex than models proposed only 10 years before, which indicates the rapid development of efficient algebraic and numeric models for these types of models. It is expected that this trend will continue.

In fact, readily scaleable, well-structured dynamic models are arguably the only hope if we are to understand the interactions

within bigger systems, such as food webs. It appears that the larger such models become, the more they will benefit from a canonical structure that allows standardized analyses and controlled comparisons. Whereas *ad hoc* models become structurally and numerically unfathomable, yielding results whose accuracy and reliability are extremely difficult to assess, the same types of analyses apply to small as to large canonical models. For instance, the same structural relationships exist between local features (kinetic orders) and global features (sensitivities), no matter how large the model. The resulting capability of canonical analyses to pinpoint problem areas in large models (72) is a tremendous asset in the diagnostics and refinement of models. It is astonishing to observe how many mathematical models are proposed without any analysis of sensitivities or robustness. The reason for this lack presumably lies in the algebraic and logistic complications of obtaining sensitivities and criteria for their evaluation.

Methods of controlled comparisons allow the researcher to focus on one or a few processes while keeping the remainder of the system unchanged. This strategy of controls is a fundamental component of the scientific process, and research in biology and chemistry is almost unthinkable without using this concept. Yet controlled comparisons are still a rare exception in mathematical modeling. Canonical modeling offers unique ways of comparing alternate candidate models because their structures are the same, models are comparable on equal footing, and their comparisons can be executed in an objective fashion.

Canonical models are particularly useful when the available data for comparison are limited. The models can be constructed from the topology of processes, i.e., from arrow diagrams that show all fluxes of material and relate which system components modulate each other. Because every parameter in a canonical model has a uniquely defined meaning, educated guesses are often sufficient to set up a model and test its features in an order-of-magnitude fashion. In some cases, even the canonical model in its symbolic form, i.e., without the specification of numerical values for all parameters, yields insight in the design of natural systems. For instance, Savageau (7) compared with purely algebraic means all imaginable patterns of feedback regulation. Similarly, Hlavacek and Savageau (62–64) elucidated patterns of gene expression, often without having to assume or measure parameter values. Given the size and complexity of environmental systems, this tolerance of canonical models of ignorance of particular parameter values is of great value.

Canonical models do not exist in total isolation from other modeling efforts. It was

shown here and elsewhere that many natural laws and accepted models are in fact canonical models, either directly or upon mathematical reinterpretation or transformation. For instance, many growth laws were shown to be simple canonical models, and it was explained why that might be so (54,58). It was shown above that exposure and transport models are simplified canonical models. Although it remains to be seen whether canonical modeling can improve or aid in analyses with these models, it is noteworthy that these models are in fact special cases. This supports the validity of the canonical modeling structure on one hand and may provide new avenues of analysis on the other. The latter was indicated with the example of the linear-logistic and proportional hazards models.

Canonical models are no panacea. It is clear that there are scenarios for which a traditional model or a new ad hoc model is superior, simpler, or both. Nonetheless, canonical modeling seems to provide a good compromise in many situations, in particular, when information is sketchy and of a diverse nature. Canonical modeling often reduces mathematical complexity, and this has theoretic as well as very practical implications. On a theoretic level, the crucial question is whether one can get away with the simpler power-law approximation of a process instead of using a mathematically more complicated structure such as a polynomial, rational, or trigonometric function. If the answer is yes, the homogeneous structure of canonical models allows algebraic and numerical analyses that are otherwise difficult or impossible to achieve and provides a repertoire of tools that elucidates critical features of complex phenomena. One might even accept some inaccuracies to obtain a canonical modeling structure that offers the advantages summarized in this review.

## REFERENCES AND NOTES

- Olden K, Klein J-L. Environmental health science research and human risk assessment. *Mol Carcinog* 14:2-9 (1995).
- Garfinkel D. Computer-based modeling of biological systems which are inherently complex: problems, strategies, and methods. *Biomed Biochim Acta* 44(6):823-829 (1985).
- Koch C, Laurent G. Complexity and the nervous system. *Science* 284:96-98 (1999).
- Laszlo E. *The Systems View of the World*. New York:George Braziller, 1972.
- Goldenfeld N, Kadanoff LP. Simple lessons from complexity. *Science* 284:87-89 (1999).
- Miller GA. The magical number seven, plus or minus two. Some limits to our capacity for processing information. *Psychol Rev* 63:81-97 (1956).
- Savageau MA. *Biochemical Systems Analysis. A Study of Function and Design in Molecular Biology*. Reading, MA: Addison-Wesley, 1976.
- Savageau MA. Critique of the enzymologist's test tube. In: *Fundamentals of Medical Cell Biology*, Vol 3A (Bitter EE, ed). Greenwich, CT:JAI Press, 1992:45-108.
- Savageau MA. A kinetic formalism for integrative molecular biology: manifestation in biochemical systems theory and use in elucidating design principles for gene circuits. In: *Integrative Approaches to Molecular Biology* (Collado-Vides J, Magasanik B, Smith TF, eds). Cambridge, MA:MIT Press, 1996:115-146.
- Savageau MA. 20 years of S-systems. In: *Canonical Nonlinear Modeling. S-System Approach to Understanding Complexity* (Voit EO, ed). New York:Van Nostrand Reinhold, 1991:1-44.
- Savageau MA. Reconstructionist molecular biology. *New Biol* 3:190-197 (1991).
- Simpson GG. Meanings of reductionism. *Science* 188:836-838 (1975).
- Yates FE. Explanation in science: is there a general theory of systems? *Am J Physiol* 233(2):R169-R170 (1977).
- Kuhn TS. *The Structure of Scientific Revolutions*. Chicago, IL: The University of Chicago Press, 1962.
- Heinrich R, Rapoport SM, Rapoport TA. Metabolic regulation and mathematical models. *Prog Biophys Mol Biol* 32:1-82 (1977).
- Welch GR. On the role of organized multienzyme systems in cellular metabolism: a general synthesis. *Prog Biophys Mol Biol* 32:103-191 (1977).
- Odadi J. Physiological significance of metabolic channelling. *J Theor Biol* 152:1-22 (1991).
- Faux MC, Scott JD. Molecular glue: kinase anchoring and scaffold proteins. *Cell* 85(1):9-12 (1996).
- Weng G, Bhalla US, Iyengar R. Complexity in biological signaling systems. *Science* 284:92-96 (1999).
- Savageau MA. *Biochemical Systems Analysis. I: Some mathematical properties of the rate law for the component enzymatic reactions*. *J Theor Biol* 25:365-369 (1969).
- Savageau MA. *Biochemical Systems Analysis. II: The steady-state solutions for an n-pool system using a power-law approximation*. *J Theor Biol* 25:370-379 (1969).
- Savageau MA. *Biochemical Systems Analysis. III: Dynamic solutions using a power-law approximation*. *J Theor Biol* 26: 215-226 (1970).
- Voit EO, ed. *Canonical Nonlinear Modeling. S-System Approach to Understanding Complexity*. New York:Van Nostrand Reinhold, 1991.
- Voit EO, Knapp RG. Derivation of the linear-logistic model and Cox's proportional hazard model from a canonical system description. *Stat Med* 18:1705-1729 (1997).
- Voit EO, Schubauer-Berigan MK. The role of canonical modeling as a unifying framework for ecological and human risk assessment. In: *Risk Assessment: Logic and Measurement* (Newman MC, Strojkan CL, eds). Chelsea, MI:Ann Arbor Press, 1998:101-139.
- Voit EO. Canonical modeling: a link between environmental models and statistics. *Austrian J Stat* 27:109-121 (1998).
- Voit EO. *Computational Analysis of Biochemical Systems. A Practical Guide for Biochemists and Molecular Biologists*. Cambridge, UK:Cambridge University Press, 2000.
- Hayt WH, Kemmerly JE. *Engineering Circuit Analysis*. New York:McGraw-Hill Book Company, 1978.
- Savageau MA, Voit EO. Recasting nonlinear differential equations as S-systems: a canonical nonlinear form. *Math Biosci* 87: 83-115 (1987).
- Lotka AJ. *Elements of Physical Biology*. Baltimore:Williams and Wilkins, 1924; reprinted as *Elements of Mathematical Biology*, New York:Dover, 1956.
- Volterra V. *Leçons sur la Théorie Mathématique de la Lutte pour la Vie*. Paris:Gauthier-Villars, 1931.
- Voit EO, Savageau MA. Accuracy of alternative representations for integrated biochemical systems. *Biochemistry* 26: 6869-6880 (1987).
- Peschel M, Mende W. The Predator-Prey Model: Do We Live in a Volterra World? Berlin:Akademie-Verlag, 1986.
- Savageau MA, Voit EO, Irvine DH. Biochemical systems theory and metabolic control theory. I: Fundamental similarities and differences. *Math Biosci* 86:127-145 (1987).
- Savageau MA, Voit EO, Irvine DH. Biochemical systems theory and metabolic control theory. II: The role of summation and connectivity relationships. *Math Biosci* 86:147-169 (1987).
- Sorribas A, Savageau MA. A comparison of variant theories of intact biochemical systems. 1: Enzyme-enzyme interactions and biochemical systems theory. *Math Biosci* 94:161-193 (1989).
- Sorribas A, Savageau MA. A comparison of variant theories of intact biochemical systems. 2: Flux oriented and metabolic control theories. *Math Biosci* 94:195-238 (1989).
- Sorribas A, Savageau MA. Strategies for representing metabolic pathways within biochemical systems theory: reversible pathways. *Math Biosci* 94:239-269 (1989).
- Torres NV, Voit E, Glez-Alcon C, Rodríguez F. An indirect optimization method for biochemical systems: description of method and application to the maximization of the rate of ethanol, glycerol, and carbohydrate production in *Saccharomyces cerevisiae*. *Biotech Bioeng* 55(5):758-772 (1997).
- Hernández-Bermejo B, Fairén V. Lotka-Volterra representation of general nonlinear systems. *Math Biosci* 140:1-32 (1997).
- Savageau MA. Michaelis-Menten mechanism reconsidered: implications of fractal kinetics. *J Theor Biol* 176:115-124 (1995).
- Savageau MA. Influence of fractal kinetics on molecular recognition. *J Mol Recognit* 6:149-157 (1993).
- Savageau MA. Development of fractal kinetic theory for enzyme-catalyzed reactions and implications for the design of biochemical pathways. *Biosystems* 47:9-36 (1998).
- Needham J. *Biochemistry and Morphogenesis*. Cambridge, UK:Cambridge University Press, 1942.
- Adolph EF. Quantitative relations in the physiological constitutions of mammals. *Science* 109:579-585 (1949).
- Lindstedt SL. Allometry: body size constraints in animal design. In: *NRC: Pharmacokinetics in Risk Assessment. Drinking Water and Health, Vol 8*. Washington, DC:National Academy Press, 1987:65-79.
- Mill T. Environmental chemistry. In: *Ecological Risk Assessment* (Suter GW II, ed). Boca Raton, FL:Lewis Publishers, 1993.
- Arnold SF, Klotz DM, Collins BM, Vonier PM, Guillelte LJ Jr, McLachlan JA. Synergistic activation of estrogen receptor with combinations of environmental chemicals. *Science* 272: 1489-1491 (1996).
- White J. The allometric interpretation of the self-thinning rule. *J Theor Biol* 89:475-500 (1981).
- Voit EO. Dynamics of self-thinning plant stands. *Ann Bot* 62: 67-78 (1988).
- Kooijman SALM. A safety factor for LC<sub>50</sub> values allowing for differences in sensitivity among species. *Water Res* 21(3): 269-276 (1987).
- Hayton WL. Pharmacokinetic parameters for interspecies scaling using allometric techniques. *Health Phys* 57:159-164 (1989).
- Suter GW II. *Ecological Risk Assessment*. Boca Raton, FL:Lewis Publ., 1993:311-364.
- Savageau MA. Allometric morphogenesis of complex systems: derivation of the basic equations from first principles. *Proc Natl Acad Sci USA* 76:6023-6025 (1979).
- Voit EO. Selecting a model for integrated biomedical systems. In: *Biomedical Modeling and Simulation* (Eisenfeld J, Witten M, Levine DS, eds). Amsterdam:Elsevier Science Publishers, 1992:465-472.
- Burger J, Gochfield M. Temporal scales in ecological risk assessment. *Arch Environ Contam Toxicol* 23:484-488 (1992).
- Eakin T. Intrinsic time scaling in a survival analysis: application to biological populations. *Bull Math Biol* 56(6):1121-1141 (1994).
- Savageau MA. Growth of complex systems can be related to the properties of their underlying determinants. *Proc Natl Acad Sci USA* 76:5413-5417 (1979).
- Savageau MA. Mathematics of organizationally complex systems. *Biomed Biochim Acta* 44:839-844 (1985).
- Irvine DH, Savageau MA. Network regulation of the immune response: alternative control points for suppressor modulation of effector lymphocytes. *J Immunol* 134:2100-2116 (1985).
- Irvine DH, Savageau MA. Network regulation of the immune response: modulation of suppressor lymphocytes by alternative signals including contra-suppression. *J Immunol* 134: 2117-2130 (1985).
- Slavacek WS, Savageau MA. Subunit structure of regulator proteins influences the design of gene circuitry: analysis of perfectly coupled and completely uncoupled circuits. *J Mol Biol* 248:739-755 (1995).
- Slavacek WS, Savageau MA. Rules for coupled expression of regulator and effector genes in inducible circuits. *J Mol Biol* 255(1):121-139 (1996).
- Slavacek WS, Savageau MA. Completely uncoupled and perfectly coupled gene expression in repressible systems. *J Mol Biol* 266:538-558 (1997).
- Shiraishi F, Savageau MA. The tricarboxylic acid cycle in dictyostelium discoideum. III: Analysis of steady state and dynamic behaviour. *J Biol Chem* 267:22926-22933 (1992).
- Sorribas A, Curto R, Cascante M. Comparative characterization of the fermentation pathway of *Saccharomyces cerevisiae* using biochemical systems theory and metabolic control analysis. Model validation and dynamic behavior. *Math Biosci* 130: 71-84 (1995).
- Ni T-C, Savageau MA. Application of biochemical systems theory to metabolism in human red blood cells. *J Biol Chem* 271 (14):7927-7941 (1996).
- Curto R, Voit EO, Sorribas A, Cascante M. Mathematical models of purine metabolism in man. *Math Biosci* 151:1-49 (1998).
- NRC. *Pharmacokinetics in Risk Assessment. Drinking Water and Health, Vol 8*. Washington, DC:National Academy Press, 1987.
- Krishnan K, Clewell HJ III, Andersen ME. Physiologically based

- pharmacokinetic analyses of simple mixtures. *Environ Health Perspect* 102 (suppl 9):151–155 (1994).
71. Savageau MA. Critique of the enzymologist's test tube. In: *Fundamentals of Medical Cell Biology*, Vol 3A (Bittar EE, ed). Greenwich, CT:JAI Press, 1992;45–108.
  72. Shiraiishi F, Savageau MA. The tricarboxylic acid cycle in dictyostelium discoideum. II: Evaluation of model consistency and robustness. *J Biol Chem* 267:22919–22925 (1992).
  73. Ni T-C, Savageau MA. Model assessment and refinement using strategies from biochemical systems theory: application to metabolism in human red blood cells. *J Theor Biol* 179: 329–368 (1996).
  74. Janszen DB. PB-PK: An Interactive Physiologically-Based Pharmacokinetic Modeling and Simulation System. [Doctoral Dissertation]. Charleston, SC:Medical University of South Carolina, 1992.
  75. Moolgavkar SH, Venzon DJ. Two event models for carcinogenesis: incidence curves for childhood and adult tumors. *Math Biosci* 47:55–77 (1979).
  76. Serio G. Two-stage stochastic model for carcinogenesis with time-dependent parameters. *Stat Probabil Lett* 2:95–103 (1984).
  77. Hemond HF, Fechner FJ. *Chemical Fate and Transport in the Environment*. San Diego, CA:Academic Press, 1994.
  78. Savageau MA. Growth equations: a general equation and a survey of special cases. *Math Biosci* 48:267–278 (1980).
  79. Savageau MA. A suprasystem of probability distributions. *Biom J* 24:323–330 (1982).
  80. Voit EO. The S-distribution. A tool for approximation and classification of univariate, unimodal probability distributions. *Biom J* 34(7):855–878 (1992).
  81. Voit EO, Yu S. The S-distribution. Approximation of discrete distributions. *Biom J* 36:205–219 (1994).
  82. Yu S, Voit EO. A simple, flexible failure model. *Biomed J* 37: 585–609 (1995).
  83. Voit EO, Balthis WL, Holser RA. Hierarchical Monte Carlo modeling with S-distributions: concepts and illustrative analysis of mercury contamination in king mackerel. *Environ Int* 21: 627–635 (1995).
  84. Voit EO, Schwacke LH. Random number generation from right-skewed, symmetric, and left-skewed distribution. *Risk Anal* 20(1):59–71 (2000).
  85. Savageau MA. Optimal design of feedback control by inhibition: dynamic considerations. *J Mol Evol* 5:199–222 (1975).
  86. Lewis DC. A qualitative analysis of S-systems: Hopf bifurcations. In: *Canonical Nonlinear Modeling. S-System Approach to Understanding Complexity* (Voit EO, ed). New York:Van Nostrand Reinhold, 1991;304–344.
  87. Reder C. Metabolic control theory: a structural approach. *J Theor Biol* 135:175–201 (1988).
  88. Heinrich R, Schuster S. *The modeling of metabolic systems. Structure, control, and optimality*. *Biosystems* 47:61–77 (1998).
  89. Ryan PB. An overview of human exposure modeling. *J Exp Anal Environ Epidemiol* 1 (4): 453–474 (1991).
  90. Savageau MA. The behavior of intact biochemical control systems. *Curr Top Cell Regul* 6:63–129 (1972).
  91. Voit EO. Utility of biochemical systems theory for the analysis of metabolic effects from low-dose chemical exposure. *Risk Anal* 20(3):393–402 (2000).
  92. Rothman KJ. *Modern Epidemiology*. Boston:Little, Brown and Company, 1986.
  93. Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research*. London:Lifetime Learning Publications (Wadsworth), 1982.
  94. Cox DR. Regression models and life tables (with discussion). *J R Stat Soc B* 68: 703–707 (1972).
  95. Voit EO, Sands PJ. Modeling forest growth. I: Canonical approach. *Ecol Model* 86:51–71 (1996).
  96. Voit EO, Sands PJ. Modeling forest growth. II. Biomass partitioning in Scots pine. *Ecol Model* 86:73–89 (1996).
  97. Voit EO. Dynamic trends in distributions. *Biomed J* 38(5):587–603 (1996).
  98. Sorribas A, March J, Voit EO. Estimating age-related trends in cross-sectional studies using S-distributions. *Stat Med* 19: 697–713 (2000).